

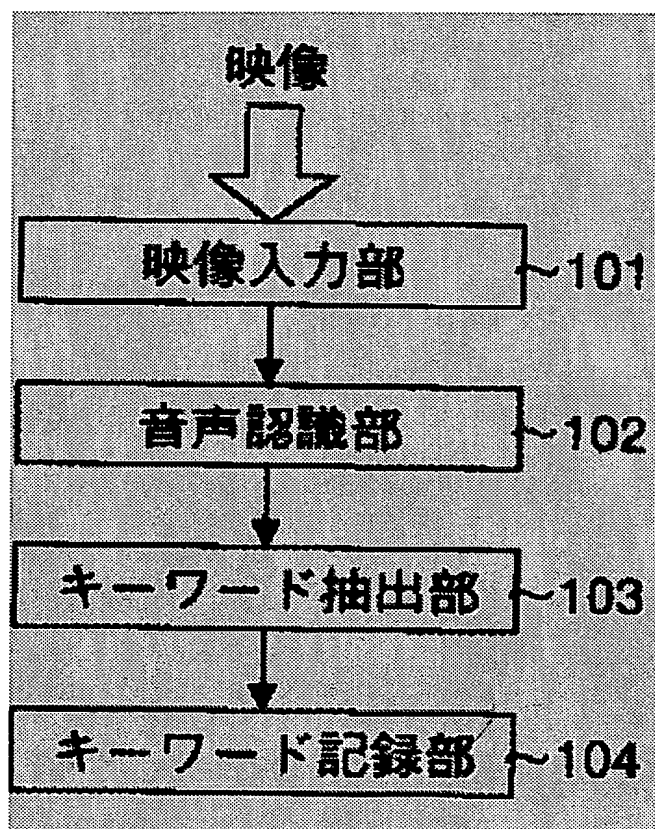
VIDEO PROCESSING APPARATUS

Patent number: JP2002171481
Publication date: 2002-06-14
Inventor: IDA HIROKO
Applicant: RICOH CO LTD
Classification:
- **International:** H04N5/91; G06F17/30; G10L15/08;
G10L15/10; G10L15/00; H04N5/93
- **European:**
Application number: JP20000369068 20001204
Priority number(s):

Abstract of JP2002171481

PROBLEM TO BE SOLVED: To make a job for attaching an index to a video image efficient so as to reduce the time and the cost required for the job.

SOLUTION: A video processing apparatus editing and retrieving a video image is provided with a video input section 101 that receives a structured video image, an audio recognition section 102 that recognizes audio in the video image, a keyword extract section 103 that extracts a keyword from the recognized audio, and a keyword recording section 104 that records the keyword together with the cross-reference with respect to the video image.



Data supplied from the esp@cenet database - Patent Abstracts of Japan

BEST AVAILABLE COPY

(19)日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11)特許出願公開番号
特開2002-171481
(P2002-171481A)

(43)公開日 平成14年6月14日(2002.6.14)

(51)IntCl ⁷	識別記号	F I	テーム(参考)
H 0 4 N 5/91		G 0 6 F 17/30	1 7 0 G 5 B 0 7 5
G 0 6 F 17/30	1 7 0		2 1 0 A 5 C 0 5 3
	2 1 0		3 4 0 B 5 D 0 1 5
	3 4 0	H 0 4 N 5/91	C
G 1 0 L 15/08		G 1 0 L 3/00	5 3 1 W

審査請求 未請求 請求項の数7 O L (全 6 頁) 最終頁に続く

(21)出願番号 特願2000-369068(P2000-369068)

(22)出願日 平成12年12月4日(2000.12.4)

(71)出願人 000006747

株式会社リコー

東京都大田区中馬込1丁目3番6号

(72)発明者 井田 裕子

東京都大田区中馬込1丁目3番6号 株式
会社リコー内

Fターム(参考) 5B075 ND16 NK31 PR04

5C053 FA14 GB11 HA30 JA05

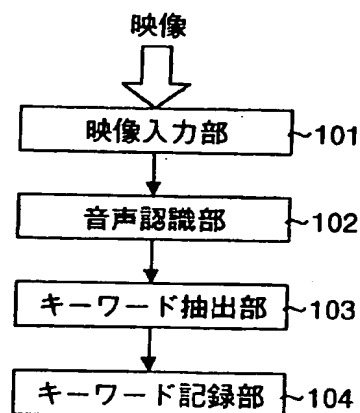
5D015 KK02 KK03 LL12

(54)【発明の名称】 映像処理装置

(57)【要約】

【課題】 映像にインデックスを付す作業を効率化し、作業にかかる時間やコストを軽減する。

【解決手段】 映像を、編集、検索する映像処理装置に対し、構造化された映像を入力する映像入力部101と、映像の音声認識する音声認識部102と、認識された音声からキーワードを抽出するキーワード抽出部103と、キーワードを映像に対する対応関係と共に記録するキーワード記録部104と、を設ける。



【特許請求の範囲】

【請求項1】 構造化された映像を、編集、検索する映像処理装置であって、
前記構造化された映像を入力する映像入力手段と、
前記映像入力手段に入力した映像の音声認識する音声認識手段と、
前記音声認識手段によって認識された音声からキーワードを抽出するキーワード抽出手段と、
前記キーワードを前記構造化された映像に対する対応関係と共に記録するキーワード記録手段と、を備えることを特徴とする映像処理装置。

【請求項2】 さらに、前記キーワード抽出手段によって抽出されたキーワードのうちのいずれかから前記構造化された映像のタイトルを選択するタイトル選択手段と、前記タイトルを前記構造化された映像に対する対応関係と共に記録するタイトル記録手段と、を備えることを特徴とする請求項1に記載の映像処理装置。

【請求項3】 前記タイトル選択手段は、前記キーワード抽出手段によって抽出されたキーワードのうちの所定の単位における出現頻度の高いキーワードをタイトルに選択することを特徴とする請求項1または2に記載の映像処理装置。

【請求項4】 さらに、キーワードの候補あるいはタイトルの候補を予め登録する候補登録手段を備え、前記構造化された映像の音声に含まれる情報のうちに前記候補登録手段に登録されているキーワードあるいはタイトルと一致する情報が含まれている場合、該情報をキーワードあるいはタイトルに選択する候補選択手段と、を備えることを特徴とする請求項1～3のいずれか一つに記載の映像処理装置。

【請求項5】 前記登録手段は、キーワードの候補あるいはタイトルの候補を選択の優先順位を付して登録することを特徴とする請求項4に記載の映像処理装置。

【請求項6】 前記キーワード記録手段が、キーワードをテキストデータまたは音声データとして記録することを特徴とする請求項1～5のいずれか一つに記載の映像処理装置。

【請求項7】 前記タイトル記録手段が、タイトルをテキストデータまたは音声データとして記録することを特徴とする請求項2～6のいずれか一つに記載の映像処理装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 本発明は、構造化された映像をさらに編集あるいは検索する映像処理装置に関する。

【0002】

【従来の技術】 近年、記録メディアに関する技術、通信技術の向上により、動画像と音声とを含むデジタルデータ（本明細書中では映像と記す）が盛んに流通するようになっている。映像の流通に伴い、記録メディアに記

録されている映像を単に時系列的に表示するのではなく、映像の必要な部分を取り出し、閲覧、編集することが可能なシステムが必要になっている。

【0003】 現在、映像の取り出し、閲覧、編集をするシステムに、映像によって表現される内容をインデックスとして記述し、インデックスの情報に基づいて映像を検索し、再構成、再編集するものがある。なお、本明細書中では、映像をインデックスが付されるグループに分類する処理を映像の構造化というものとする。構造化は、映像を階層的に分類することによってなされる場合も多い。

【0004】 映像にインデックスを付す方法には、物理的に記述する方法と言語的に記述する方法とがある。物理的記述方法は、例えばシーンの切換えやカメラアングルといった規則にしたがう内容を記述する方法である。また、言語的記述方法は、例えば映像における被写体の名称や被写体の行動などを文章で記述する方法である。

【0005】

【発明が解決しようとする課題】 言語的記述方法でインデックスを付す場合、最も簡単な方法はマニュアルによるものである。ところが、大量の映像にマニュアルで言語的記述方法によるインデックスを付す作業は作業量が大きく、作業にかかる時間やコストが膨大なものになるという欠点がある。このため、内容を反映したインデックスを、より効率的に映像に付すことができる映像処理装置が要求されている。

【0006】 本発明は、上記の点に鑑みてなされたものであり、映像にインデックスを付す作業を効率化し、作業にかかる時間やコストを軽減することができる映像処理装置を提供することを目的とする。

【0007】

【課題を解決するための手段】 上記した課題を解決し、目的を達成するため、請求項1に記載の発明にかかる映像処理装置は、構造化された映像を、編集、検索する映像処理装置であって、前記構造化された映像を入力する映像入力手段と、前記映像入力手段に入力した映像の音声認識する音声認識手段と、前記音声認識手段によって認識された音声からキーワードを抽出するキーワード抽出手段と、前記キーワードを前記構造化された映像に対する対応関係と共に記録するキーワード記録手段と、を備えることを特徴とする。

【0008】 この請求項1に記載の発明によれば、構造化された映像に対し、映像に含まれる音声に基づいて自動的にキーワードを付すことができる。このため、キーワードをマニュアル（作業者が映像を見てキーワードを付す作業）で付すことに比べて作業者にかかる負荷を大幅に軽減し、作業時間をも短縮することができる。

【0009】 請求項2に記載の発明にかかる映像処理装置は、さらに、前記キーワード抽出手段によって抽出されたキーワードのうちのいずれかから前記構造化された

映像のタイトルを選択するタイトル選択手段と、前記タイトルを前記構造化された映像に対する対応関係と共に記録するタイトル記録手段と、を備えることを特徴とする。

【0010】この請求項2に記載の発明によれば、構造化された映像に対し、キーワードに基づいて自動的にタイトルを付すことができる。

【0011】請求項3に記載の発明にかかる映像処理装置は、前記タイトル選択手段は、前記キーワード抽出手段によって抽出されたキーワードのうちの所定の単位に
10 における出現頻度の高いキーワードをタイトルに選択することを特徴とする。

【0012】この請求項3に記載の発明によれば、より多くのキーワードを映像のタイトルに反映させることができる。

【0013】請求項4に記載の発明にかかる映像処理装置は、さらに、キーワードの候補あるいはタイトルの候補を予め登録する候補登録手段を備え、前記構造化された映像の音声に含まれる情報のうちに前記候補登録手段に登録されているキーワードあるいはタイトルと一致する
20 情報が含まれている場合、該情報をキーワードあるいはタイトルに選択する候補選択手段と、を備えることを特徴とする。

【0014】この請求項4に記載の発明によれば、予め登録されたキーワードやタイトルの候補からキーワードやタイトルを選択することができる。

【0015】請求項5に記載の発明にかかる映像処理装置は、前記登録手段が、キーワードの候補あるいはタイトルの候補を選択の優先順位を付して登録することを特徴とする。

【0016】この請求項5に記載の発明によれば、登録手段にタイトル候補語句を選択の優先順位を付して登録しておくことができる。

【0017】請求項6に記載の発明にかかる映像処理装置は、前記キーワード記録手段が、キーワードをテキストデータまたは音声データとして記録することを特徴とする。

【0018】この請求項6に記載の発明によれば、キーワード記録手段が、キーワードをテキストデータ（語句）で記録する他、音声のデータとして記録することも
40 できる。

【0019】請求項7に記載の発明にかかる映像処理装置は、前記タイトル記録手段が、タイトルをテキストデータまたは音声データとして記録することを特徴とする。

【0020】この請求項7に記載の発明によれば、タイトル記録手段が、タイトルをテキストデータ（語句）で記録する他、音声のデータとして記録することもできる。

【0021】

【発明の実施の形態】以下に添付図面を参照して、この発明にかかる映像処理装置の好適な実施の形態1～4を詳細に説明する。

（実施の形態1）図1は、本発明の実施の形態1の映像処理装置の構成を説明するためのブロック図である。図示した構成は、構造化された映像を、編集、検索する映像処理装置である。そして、構造化された映像を入力する映像入力部101と、入力した映像の音声を認識する音声認識部102と、認識された音声からキーワードを抽出するキーワード抽出部103と、キーワードを構造化された映像に対する対応関係と共に記録するキーワード記録部104と、を備えている。

【0022】構造化された映像は、映像入力部101から本実施の形態の映像処理装置に入力する。構造化された映像は複数のセグメントで構成されている。音声認識部102は、音声認識技術を用いて映像に含まれる音声を認識する。キーワード抽出部103は、認識された音声から各セグメントごとにキーワードを抽出する。

【0023】キーワードの抽出は、例えば、後述するように、予めキーワードの候補を映像処理装置に記録しておき、音声に含まれる語句に記録されたキーワードの候補と一致するものがあった場合、この候補と一致する語句をキーワードとして抽出してもよい。また、例えば、音声において認識された回数の多い語句をキーワードとしてもよい。

【0024】上記した実施の形態1によれば、構造化された映像を構成するセグメントに対し、映像に含まれる音声に基づいて自動的にキーワードを付すことができる。このような実施の形態1は、キーワードをマニュアル（作業者が映像を見てキーワードを付す作業）で付すことに比べて作業者にかかる負荷を大幅に軽減し、作業時間をも短縮することができる。

【0025】（実施の形態2）次に、実施の形態2の映像処理装置について説明する。図2は、実施の形態2の映像処理装置を説明するためのブロック図である。なお、図2に示した構成のうち、図1に示した構成と同様の構成については同様の符号を付して説明を一部省くものとする。

【0026】実施の形態2の映像処理装置は、さらに、キーワード抽出部103によって抽出されたキーワードのうちのいずれかから構造化された映像のタイトルを選択するタイトル選択部201と、選択されたタイトルを構造化された映像に対する対応関係と共に記録するタイトル記録部202と、を備えたものである。

【0027】上記した映像処理装置において、映像入力部101に映像と共に入力し、音声認識部102で認識された音声は、キーワード抽出部103でキーワードを抽出される。抽出されたキーワードは、タイトル選択部201に入力する。タイトル選択部201は、抽出されたキーワードの中からタイトルを選択する。
50

5

【0028】タイトルの選択は、例えば、抽出されたキーワードのうちの抽出頻度が高いものを選択することによって行うことができる。キーワードの抽出頻度に基づいてタイトルを選択する場合、図3に示すように、図2の構成にさらに出現頻度算出部を設けてキーワードの抽出頻度を算出するようにしてもよい。また、後述するように、予めタイトルの候補を映像処理装置に記録しておき、キーワードに記録されたタイトルの候補と一致するものがあった場合、この候補と一致するキーワードをタイトルとして抽出してもよい。

【0029】上記した実施の形態2によれば、構造化された映像に対し、キーワードに基づいて自動的にタイトルを付すことができる。実施の形態2は、タイトルをマニュアル（作業者が映像を見てタイトルを付す作業）で付すことに比べて作業者にかかる負荷を大幅に軽減し、作業時間をも短縮することができる。

【0030】また、実施の形態2によれば、より多くのセグメントのキーワードを映像のタイトルに反映させることができる。このため、映像を構成するセグメントの内容を反映したタイトルを映像に付すことができる。

【0031】（実施の形態3）次に、実施の形態3の映像処理装置について説明する。図4は、実施の形態3の映像処理装置を説明するためのブロック図である。なお、図4に示した構成のうち、図1に示した構成と同様の構成については同様の符号を付して説明を一部省くものとする。

【0032】実施の形態3の映像処理装置は、さらに、キーワードの候補（キーワード候補語句）あるいはタイトルの候補（タイトル候補語句）が予め登録された登録部403およびキーワード・タイトル照合部401、キーワード・タイトル記録部402を備えている。キーワード・タイトル照合部401は、構造化された映像の音声に含まれる語句のうちにキーワード語句、あるいはタイトル候補語句と一致する語句が含まれている場合、キーワード語句に一致した語句をキーワードとして抽出する。また、タイトル語句に一致した語句をタイトルとして抽出する（タイトルは、抽出されたキーワードのうちのタイトル語句と一致するものとしてもよい）。

【0033】上記した映像処理装置において、映像入力部101に映像と共に入力し、音声認識部102で認識された音声は、キーワード・タイトル照合部401において登録部403に登録されているキーワードと照合する。そして、音声に登録部403のキーワード候補語句と一致する語句があった場合、一致した語句を、語句が含まれるセグメントのキーワードとして抽出する。

【0034】なお、実施の形態3は、抽出されたキーワードのうちからタイトルを抽出するものとする。このため、キーワード・タイトル照合部401は、抽出したキーワードのうちにタイトル候補語句に一致するキーワードがある場合、このキーワードをタイトルとする。

6

【0035】キーワード・タイトル照合部401によって抽出されたキーワードおよびタイトルは、キーワード・タイトル記録部402に記録される。なお、本発明の映像処理装置は、音声認識部102によって認識された音声にタイトル候補語句と一致する語句があった場合、この語句がキーワードとして抽出されていない場合にもタイトルとすることもできる。

【0036】上記した実施の形態3によれば、予め登録されたキーワードやタイトルの候補からキーワードやタイトルを選択することにより、オペレータ側の意図をキーワードやタイトルに反映することができる。

【0037】また、実施の形態3は、登録部403にキーワード候補語句、タイトル候補語句を選択の優先順位を付して登録しておくことができる。このようにすれば、よりオペレータ側の意図をキーワードやタイトルに反映することができる。

【0038】さらに、実施の形態1～3は、キーワード記録部104、タイトル記録部202、キーワード・タイトル記録部402が、キーワードやタイトルをテキストデータ（語句）で記録する他、音声のデータとして記録することもできる。このようにすれば、キーワードやタイトルを、語句の他、音声として映像に付すことができる。

【0039】

【発明の効果】以上説明したように、請求項1に記載の発明は、キーワードをマニュアル（作業者が映像を見てキーワードを付す作業）で付すことに比べて作業者にかかる負荷を大幅に軽減し、作業時間をも短縮することができるという効果を奏する。このような請求項1に記載の発明は、映像にインデックスを付す作業を効率化し、作業にかかる時間やコストを軽減することができるものといえる。

【0040】請求項2に記載の発明は、タイトルをマニュアル（作業者が映像を見てタイトルを付す作業）で付すことに比べて作業者にかかる負荷を大幅に軽減し、作業時間をも短縮することができるという効果を奏する。このような請求項2に記載の発明は、映像にインデックスを付す作業を効率化し、作業にかかる時間やコストを軽減することができるものといえる。

【0041】請求項3に記載の発明は、映像に付されたキーワードを反映したタイトルを映像に付すことができるという効果を奏する。

【0042】請求項4に記載の発明は、オペレータ側の意図をキーワードやタイトルに反映することができる。

【0043】請求項5に記載の発明は、よりオペレータ側の意図をキーワードやタイトルに反映することができる。

【0044】請求項6に記載の発明は、キーワードを、語句の他、音声として映像に付すことができる。

【0045】請求項7に記載の発明は、タイトルを、語句の他、音声として映像に付することができる。

【図面の簡単な説明】

【図1】本発明の実施の形態1の映像処理装置の構成を説明するためのブロック図である。

【図2】本発明の実施の形態2の映像処理装置の構成を説明するためのブロック図である。

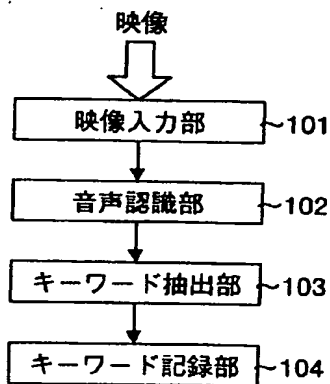
【図3】本発明の実施の形態2の他の映像処理装置の構成を説明するためのブロック図である。

【図4】本発明の実施の形態3の映像処理装置の構成を説明するためのブロック図である。

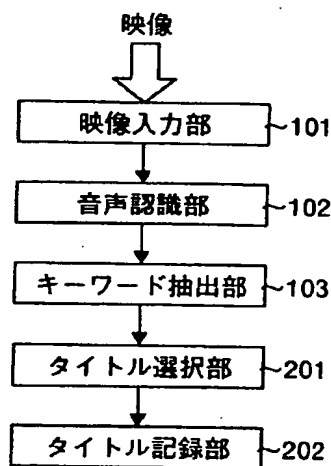
* 【符号の説明】

- 101 映像入力部
- 102 音声認識部
- 103 キーワード抽出部
- 104 キーワード記録部
- 201 タイトル選択部
- 202 タイトル記録部
- 401 キーワード・タイトル照合部
- 402 キーワード・タイトル記録部
- 403 登録部

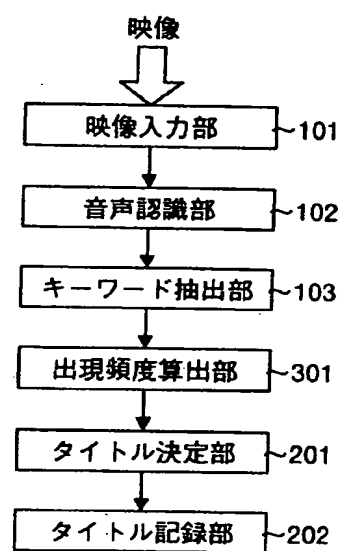
【図1】



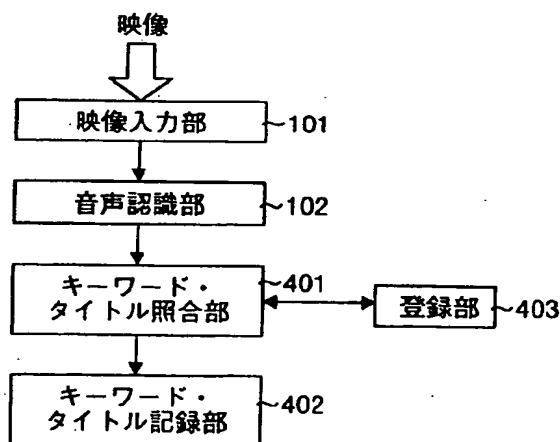
【図2】



【図3】



【図4】



フロントページの続き

(51)Int.Cl.⁷

識別記号

FI

テーマコード (参考)

G10L 15/10

G10L 3/00

551G

15/00

H04N 5/91

N

H04N 5/93

5/93

Z